Stochastic Process Model of Mortality and Aging: Application to Longitudinal Data

I.Akushevich[1], L.Akushevich[1], K.Manton[1], A.Yashin[2]

[1] *Center for Demographic Studies, Duke University, 2117 Campus Drive, Box 90408, Durham, NC 27708-0408, USA*
E-mail: aku@cds.duke.edu

[2] *Max-Planck Institute for Demographic Research, Doberaner Str. 114, 18057 Rostock, GERMANY*

(Received 22 August 2002)

Two models for making population forecasts are presented and discussed. One of them is based on an analytical Gaussian solution of the Kolmogorov-Fokker-Planck equation and regression model for risk factor dynamics. The second is based on more general diffusion type stochastic differential equations. Numerical schemes for parameter estimation from longitudinal data and for forecasting are discussed for both.

Key words: Random walk model, Kolmogorov-Fokker-Planck equation, stochastic processes, demography, health forecasting

PACS numbers: 05.10.Gg

1 Introduction. Aims and definitions

Despite numerous experimental and theoretical attempts to understand the forces shaping the age-specific human mortality curve at late ages many problems remain unsolved. Numerous explanations based on the concept of heterogeneity in mortality, individual adaptation, hormesis or on some general properties of complex multistage dynamic systems sometimes contradict each other (see [1] and [2] for review). Using only population survival data restricts analysis to simple mortality (time to failure) models.

Additionally, theoretical insights on human failure processes runs into difficulty based on new findings from experimental and population studies. The Gompertz hazard generally does not fit well in human populations at later ages ([3, 4, 5]). Even models which assume that the aging parameters are specific to individuals sometimes runs into difficulty when attempting to interpret parameter values as biologically meaningful, e.g., use of a gamma or inverse Gaussian mixed Gompertz ([6]). For example, longitudinal studies clearly show that the risk heterogeneity of a population is altered over time by behavioral factors and environmental exposures. Thus, while a continuously mixed mortality model is an empirical enhancement over the standard Gompertz when explaining the distribution of time to death it is clearly only an improved approximation when individual physiological states are clearly observed to change over time. Also relevant are recent observation on the fundamental processes of senescence. For example, an analysis based on the logic of the thermodynamics of protein denaturation leading to organism death leads to a Weibull hazard - and very different estimates of physiological interpretations of model parameters (e.g., [7, 8]). For a long time a dominant view on cellular limits to human longevity were based on the arguments of Hayflick ([9]), Martin ([10]) found that there were factors influencing those processes in that cells cultured from persons aged 30 to 80 show a much slower process of senescence than suggested by observed human longevity. Cristofalo et al. ([11]) went further and found that, in healthy individuals there was no apparent loss of replicability in some cell

lines. This suggested that it was invalid to view human longevity as a static property of a physically closed system. This suggested the need for stochastic state space models estimated from longitudinally followed population. It was also clear that for population forecasts of health to be accurate both the internal structures of humans, and exogenous inputs to those systems, had to be better described.

More data are needed to estimate theoretically satisfactory models of human aging and mortality. Many existing longitudinal and epidemiological studies and national longitudinal surveys contain information necessary for a better understanding of the regularities of state dependent mortality dynamics. By measuring age-related changes in health and physiological covariates together with life span these studies open up new opportunities for analyzing the dynamics of aging process and mortality with appropriate models.

Unfortunately those data are often under analyzed which has led to the generation of erroneous conclusions about human health and how to improve it. Use of the multiple logistic and Cox regression model do not fully utilize the information in the stochastic dynamics of the risk factors. For example, models looking at relative risks understate often the absolute importance of risk factors at advanced ages. More important is that the dynamics and interactions of risk factors are often erroneously described. For example, in analyses of the Framingham Heart Study data with a quadratic hazard it was not possible to exclude "optimal" values of serum cholesterol up to 210 or 220 mg/dl. This is consistent with the observations that the science of blood lipids is more complex than acknowledged in the original epidemiological analyses with many lipid subtypes, including healthy types, identified. Recent observations have suggested many dietary recommendation surrounding fat and carbohydrate intake needed to be complete re-evaluated and that many other risk factors, such as homocysteine ([12, 13]), are important. Also important for state space models is that there appeared to be a small number of fundamental processes underlying many disease states, e.g., lipid metabolism for circulatory disease, cancer and Alzheimer's; oxidative mechanisms for carcinogenesis, heart disease and Alzheimer's. As a consequence, we propose the use of a state space model based on an individual random walk.

In this report we improve the classical random walk model (RWM) [14, 15, 16] to make practical predictions and forecast using longitudinal data on human health. Developing a useful software package for such effects has to deal with:

- analysis of longitudinal data bases;
- construction of models describing these data bases;
- developing a numerical scheme for parameter estimation;
- forecasting with this model using estimated parameters.

In Section 2 of the report we make short survey of the mathematical formalisms of the RWM and rewrite formulas in a form suitable for numerical analysis. We describe the Framingham Heart Study data base and subsets of these data which are used in our analysis. Numerical results based on the formulas are in Section 3. Problems with missing data, and its central role in analyses, are discussed there. An alternate approach based on stochastic process models (see review [17]) is considered in Section 4. Generalizations to include non-linear dynamic and chaotic effects are discussed in Section 5. Final comments and conclusions are in Section 6.

2 Random walk model

It is possible to model the time to failure (death) of an organism using a random walk model with "manholes". The stochastic differential equations for this random walk are for state dynamics,

$$dx_w(t) = u(x_w, t)dt + d\xi(x_w, t),$$

and the state dependent distribution of "manholes",

$$dP(x_w) = -\mu(x_w, t)P(x_w)dt$$

First equation represents changes for organism w on each of J coordinate dimensions, $x = (x_j, j = 1, 2, ..., J)$ at time t. Movement $dx_w(t)$ of individual w during a time dt is the sum of deterministic $u(x_w, t)dt$ and random walk $d\xi(x_w, t)$ contributions. $P(x_w)$ is the probability of surviving for w at the point x_w for time t. The change of survival probability $dP(x_w)$ is proportional to the probability of mortality $\mu(x_w, t)dt$. The meaning of parameters, and state space dimensionality, are usually defined by experimental longitudinal data of parameters used in analyses.

The Framingham Heart Study is an example of a longitudinal data base, where detailed medical characteristics of individuals are measured at approximately equal time intervals between measurements. The original Framingham Heart Study cohort consisted of respondents of a random sample of 2/3 of adults, 30 to 62 years of age, residing in Framingham, Massachusetts in 1948. Of the original 5209 persons, there are approximately 1095 known alive as of February 1998. There are, in the data set available to us, 23 two-year follow-ups 1948 to 1998. We used 11 risk factors [18]: age, pulse pressure, diastolic blood pressure, serum cholesterol, blood glucose, hematocrit, vital capacity index, smoking, left ventricular hypertrophy, pulse rate.

It is possible to reformulate the RWM model in terms of the Kolmogorov-Fokker-Planck (KFP) equation [19]

$$\frac{\partial f}{\partial t} = -\sum_j u_j \frac{\partial f}{\partial x_j} - f\sum_j \frac{\partial u_j}{\partial x_j} + \frac{1}{2}\sum_i \sum_j \sigma^0_{ij} \frac{\partial^2 f}{\partial x_j x_j} - \mu f$$

In this form the task of describing collective population movement over the state space is reduced to a task of describing one person making random walks averaged over other individuals in the population. The first term corresponds to the "drift" u of mean values during these random walks. This drift can be x-dependent – that is reflected in the second term. The third term takes into account diffusion σ^0. The last term corresponds to mortality. It changes the normalization of the multivariate distribution function over time.

One way to solve the KFP equation is to make additional assumptions which reduce to the assumption that the population distribution can be described at time t as a multivariate normal distribution $N(l_t, \nu_t, V_t)$, whose three parameters represent the population size (l_t), the vector of physiological variable means (ν_t), and the variance-covariance matrix (V_t). We also need to make assumptions about the linear model for the dynamic risk factor variables over time. Here two possible models are considered. The first is based on the assumption that only the current year information influences the risk factors of next year exam. The two-year model takes into account memory effect, next exam risk factors are defined by two previous exams. The basic model assumes that initial value of vector of physiological variables is normally distributed, the dynamics are linear and the hazard function is quadratic:

$$\mu = \left(\mu_0 + bx + \frac{1}{2}x^T Bx\right)e^{\theta t} \qquad (1)$$

where model parameters are μ_0, b (J-vector) and B ($J \times J$-matrix); $e^{\theta t}$ is an exponential term reflecting unmeasured state variables correlated with age.

Regression formula for one-year model is

$$x_{t+1} = u_0 + Rx_t + \epsilon, \qquad (2)$$

where vector u_0 and matrix R are model parameters and ϵ is a vector of residials. The following formulas are used to adjust the distribution for mortality:

$$\begin{aligned} \nu^*_t &= \nu_t - V^*_t(b_t + B_t\nu_t) \\ V^*_t &= (V_t^{-1} + B_t)^{-1}. \end{aligned}$$

And also

$$V_{t+1} = \Sigma + RV^*_t R^T,$$

where Σ is variance-covariance matrix of residuals. Hazard parameters b, B in these formulas are included in the definition of the mortality model [20].

Similarly for two-year model we have

$$\begin{aligned} x_{t+1} &= u_0 + Rx_t + Tx_{t-1} + \epsilon, \\ V_{t+1} &= \Sigma + RV^*_t R^T + TV^{**}_{t-1}T^T. \end{aligned}$$

Quantities with two stars are characteristics of normal distributions for individuals surviving two time periods:

$$\begin{aligned} \nu_t^{**} &= \nu_t^* - V_t^{**}(b_{t-1} + B_{t-1}\nu_t^*), \\ V_t^{**} &= ({V_t^*}^{-1} + B_{t-1})^{-1}. \end{aligned}$$

In the both models, values of age should be fixed. In this case the Gaussian distribution will be conditional on age. It amounts to changing matrices V and Σ to ($k = age$)

$$\begin{aligned} V_{ij}^c &= V_{ij} - \frac{V_{ik}V_{kj}}{V_{kk}}, \\ \Sigma_{ij}^c &= \Sigma_{ij} - \frac{\Sigma_{ik}\Sigma_{kj}}{\Sigma_{kk}} \end{aligned}$$

Hazard parameters β_{ij} in the quadratic function are estimated with maximum likelihood procedures. Vector b, μ_0 and matrix B (see Eq. (1)) are related to the hazard coefficients as,

$$\mu_0 = \beta_{00}^2 \qquad b_i = \beta_{00}\beta_{0i}, \quad B_{ik} = \sum_{j=0}^{K} \beta_{jk}\beta_{ji}.$$

where $K = \min(i, j, i_r - 1)$. i_r defines a number of parameters β_{jk}. We take $i_r = 4$ for Framingham data.

The (log) likelihood can be expressed in the following form

$$\mathcal{L} = -\mathcal{L}_r + \sum_{i=1}^{I} \ln(e^{\mu_i} - 1)$$

where i runs over individuals ($i = 1, 2, ..., I$),

$$\mu_i = \sum_{j=0}^{i_r-1} \sum_{l=j}^{J} \sum_{m=j}^{J} \beta_{jl}\beta_{jm}\xi_{li}\xi_{mi},$$

$$\mathcal{L}_r = \sum_{j=0}^{i_r-1} \sum_{l=j}^{J} \sum_{m=j}^{J} \beta_{jl}\beta_{jm}r_{lm},$$

$$r_{lm} = \sum_{i=1}^{I} \bar{x}_{li}\bar{x}_{mi}$$

Here $\bar{x}_{ji} = Ax_{ji}$ and A is a age-dependent factor defined by a model. $\xi_{li} = \bar{x}_{ji}$ for last survey before death and $\xi_{li} = 0$ otherwise (see [20] for details).

3 Numerical results

3.1 Missing data

The database has extensive missing data. Additional assumptions are required to fill those values before estimating the model. One way is to use the following procedure: if a value between measurements is missed then it is calculated as the average proportional (for example, 20-missed-missed-80 produces 20-40-60-80); however if there are data only from one side (20-missed-missed) then missing values are set equal to the nearest measure: (20-20-20). This is so called "simple mean value" method of filling missing data.

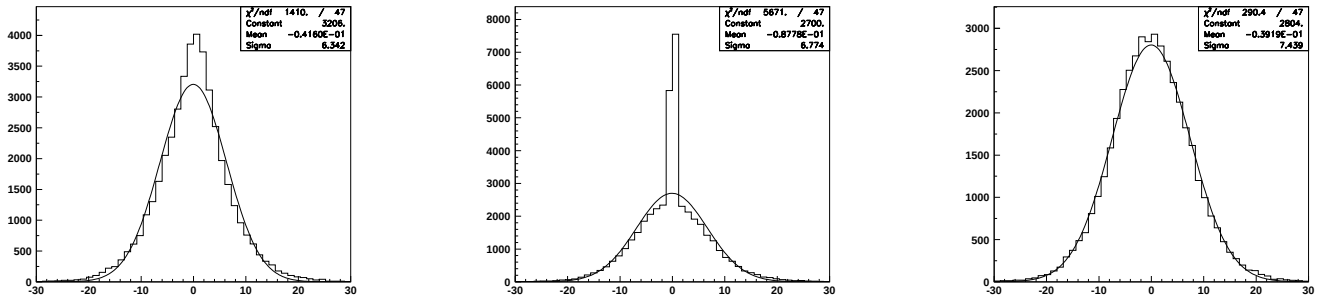
Another method - "deterministic" - at the first stage uses the results of the procedure above. It allows us to calculate an initial regression model for the data. Predictions of the model are used to fill in missing data. We truncate predicted values for data points with extreme residuals with deviation 3σ. Since changing the data changes the regression coefficients, this procedure is repeated until there are no coefficient changes for the current iteration. This procedure makes the generated database model-dependent.

Residuals for one of the risk factor variables are in Figure 1. The distribution is Gaussian with a large peak for $\epsilon = 0$. This is a result of our procedure to fill missing data, when all missing data were changed to their predicted values with $\epsilon = 0$. We see that the procedure is not statistical but deterministic.

One solution to this problem is to use Monte-Carlo simulation of deviations from theoretically predicted values generated in accordance with a Gaussian distribution with observed σ. In this case and for one-year model missing data are simulated using,

$$x_{t+1} = u_0 + Rx_t + \epsilon$$

with ϵ's simulated as a Gaussian distributed with observed σ. In this case ϵ's have the correct distribution and this peak disappears. One theoretical issue to evaluate is whether the innovation term ϵ is viewed or generated by sampling or whether it is viewed or generated by stochastic factors.



a) Mean b) Deterministic c) Monte-Carlo

FIG. 1. Residuals for x_3 diastolic blood pressure within different approaches for filling missing data.

Using formulas from the previous section one can predict behavior of risk factor versus age for starting data. Results for projections of some risk factors and life expectancy are presented in Figures 2, 3 and 4.

4 An approach based on diffusion type stochastic differential equations

An alternate approach to make projections is described in [17]. It has several advantages compared to classical RWM. All parameters are defined within a joint likelihood function, so only one numerical procedure is needed to estimate both dynamic and mortality parameters. Within this approach it is not necessary to use ancillary procedures to fill in missing data. The projection is obtained as an explicit solution of differential equations, so the times between surveys are not fixed. Therefore, it is possible to make predictions in a 'year by year' or 'month by month' scheme.

4.1 The model

Similar to the Random Walk Model described above, this model is defined by a system of stochastic differential equations. It is assumed, as above, that mortality is a quadratic function of $x(t)$

$$\mu(x(t),t) = \mu_0(t) + 2b(t)x(t) + x^*(t)B(t)x(t)$$

$$dx(t) = (a_0(t) + a_1(t)x(t))dt + a_2(t)dW_t$$

For a Gaussian process the likelihood can be written,

$$L = \prod_{i=1}^{N} \hat{\mu}(\tau_i, \hat{x}(\tau_i))^{\delta_i} \exp\left(-\int_0^{\tau_i} du\hat{\mu}(u, \hat{x}_i(u))\right) \times \prod_{j=1}^{k_i} f(x_i(t_j)|\hat{x}_i(t_{j-1})); \quad (3)$$

where $f(x_i(t_j)|\hat{x}_i(t_{j-1}))$ is a Gaussian density distribution conditional on prior observations, τ_i are ages of death, $delta_i$ are indicators of sensoring, t_j are observation times, $\hat{x}_i(t_j)$ are discrete time observations. Indexes i and j run a.) over individual in data base and b.) exams of each individual, respectively. The equation,

$$\hat{\mu}(\hat{x}(t),t) = m^*(t)B(t)m(t) + 2b(t)m(t) + \mathrm{tr}(B(t)\gamma(t)) + \mu_0(t),$$

has the sense of a right-continuous mortality rate.

Vector $m(t)$ and matrix $\gamma(t)$ are defined by system of ordinary differential equations at intervals $[t_j, t_{j+1})$

$$\begin{aligned} \frac{dm(t)}{dt} &= a_0(t) + (a_1(t) - 2b(t))m(t) - 2\gamma(t)B(t)m(t) \\ \frac{d\gamma(t)}{dt} &= a_1(t)\gamma(t) + \gamma(t)a_1^*(t) + a_2(t)a_2^*(t) - 2\gamma(t)B(t)\gamma(t). \end{aligned} \quad (4)$$

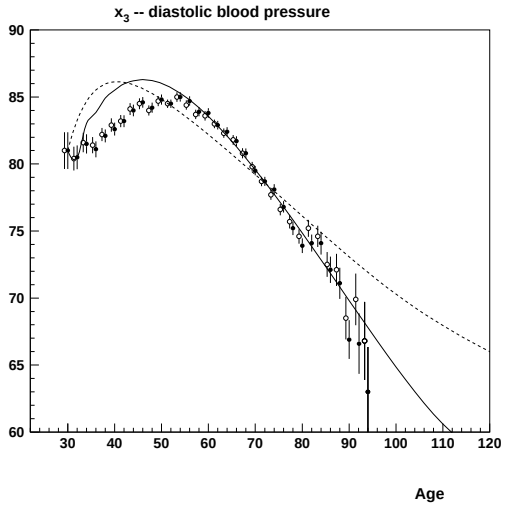


FIG. 2. Data and predictions for one-year (open circles, dashed line) and two-year (full circles, solid line) models for diastolic blood pressure (x_3)

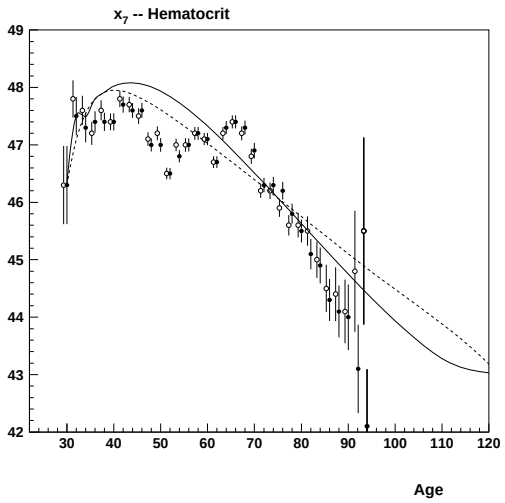


FIG. 3. Data and predictions for one-year(open circles, dashed line) and two-year (full circles, solid line) models for hematocrit (x_7)

FIG. 4. Predictions for one-year (dashed line) and two-year (solid line) models for life expectancy

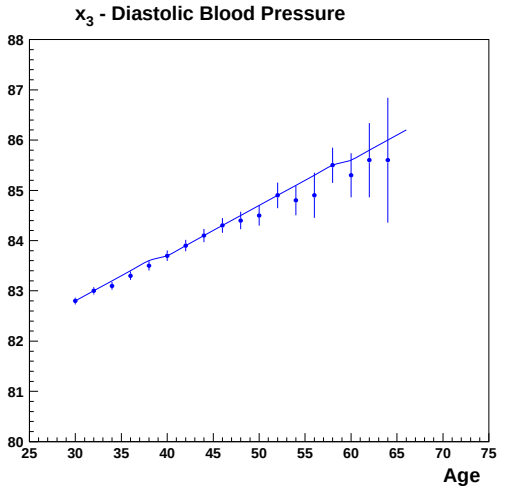


FIG. 5. Simulated data (circles) and projection (line) for diastolic blood pressure

with bound conditions $m(t_j) = \hat{x}(t_j)$, $\gamma(t_j) = 0$. Specific assumptions about the explicit form of time dependence of the functions $a_{0,1,2}(t)$, $b(t)$ and $B(t)$ have to be made. For example, it is reasonable to assume a Gompertz dependence for the hazard parameters.

4.2 Life Time Generator

Our first test is the analysis of simulated data and extraction of coefficients, which were used for simulation. Set of stochastic differential equations and assumptions about a Gaussian distribution give us the opportunity to develop a simulation strategy.

Constructed in this way the generator can be associated with a generator of life, because it reproduces all features of a cohort life span.

The strategy can be defined in several steps. First the theoretical cohort (i.e. 100,000 individuals with normally distributed risk factors) is simulated. For simplicity all individuals can be taken at the same age. However the initial risk factor values have to be normally distributed. Mortality and survival functions are calculated for each individual at each time interval. It allows us to define randomly if an individual survives in the time interval or not. For individuals surviving for time spot τ, risk factor values are simulated for next age period as,

$$x(t+\tau) = x(t) + D_1\tau + \sqrt{D_2\tau}\xi_r$$

$$D_1 = a_0(t) + a_1(t)x(t), \qquad D_2 = a_2(t)a_2^*(t)$$

This procedure can solve many tasks for life table construction as well as provide numerical and analytical tests of our computational procedures.

4.3 Numerical and analytical tests

A simulated data base can be used for parameter estimations. As a result of maximization of the likelihood one obtains parameters for the functions ($a_{0,1,2}(t)$, $\mu_0(t)$, $b(t)$ and $B(t)$) for simulation. Results for these coefficients can be used for projections (see Figure 5).

Analytic tests can be done for the simplified ($b = 0$, $B = 0$, $a_1 = 0$) one-dimensional model:

$$\mu = \mu_0 e^{\theta t}, \quad dx = a_0 dt + a_2 dW$$

The system of differential equations (4) can be solved analytically,

$$m(t) = m(t_0) + (t - t_0)a_0 \quad \gamma(t) = (t - t_0)a_2^2,$$

Integration in (3) can be also done analytically,

$$\int_0^{\tau_i} dt\ \mu_0\ e^{\theta t} = \frac{\mu_0}{\theta}(e^{\theta\tau_i} - 1)$$

The likelihood function (3) can be split into two terms

$$\begin{aligned} L &= \sum_i [\ln(\mu_0 e^{\theta\tau_i}) - \frac{\mu_0}{\theta}(e^{\theta\tau_i} - 1)] \\ &\quad -\frac{1}{2}\sum_{ij}\left[\ln\gamma + \frac{(\Delta x - a_0 y)^2}{\gamma}\right] \end{aligned} \quad (5)$$

where y is the time interval between simulated risk factor values for an individual and $\Delta x = x^{next} - x$ is simulated difference of risk factors; τ_i is the failure time. We assume that all individuals in the cohort finally die.

The maximization of (5) can be performed analytically. The first term gives

$$\mu_0 = \frac{I\theta}{\sum_i \exp(\theta\tau_i) - 1}. \quad (6)$$

The second leads to,

$$\begin{aligned} a_0 &= \frac{\sum \Delta x}{yIJ} \\ a_2^2 &= \frac{1}{\Delta t IJ}\left(\sum(\Delta x)^2 - (\sum \Delta x)^2\right). \end{aligned} \quad (7)$$

I and J are numbers of individuals and risk factors, respectively.

Equations (6) and (7) give estimates of parameters in terms of sums of the simulated quantities: τ_i and Δx. Sums can be calculated analytically for simulated data using simulation (see Section 4.2). The estimates must give exactly our initial parameters μ_0 and $a_{0,2}$. If so, it provides an analytical test of the model. The following calculation demonstrates this.

The cohort population function $l(t)$, being the solution of the differential equation $dl/dt = -\mu l$, has the form,

$$l(t) = l_0 \exp\left(-\frac{\mu_0}{\theta}(e^{\theta t} - 1)\right)$$

Parameter μ_0 in (6) can be written as $\mu_0 = \theta/\langle g(t)\rangle$, with $g(t) = \exp(\theta t) - 1$. Averaging in accordance with

$$\frac{1}{I}\sum_i [e^{\theta\tau_i} - 1] = \langle e^{\theta t} - 1\rangle = \frac{1}{l_0}\int_0^\infty d[e^{\theta t} - 1]\ l(t) = \frac{\theta}{\mu_0}.$$

That confirms the estimate of μ_0 (6).

Sums in (7) can be estimated from the generation which is reduced for our case,

$$\Delta x = a_0 \Delta t + a_2 \sqrt{y} \xi_r,$$

where ξ_r are standardized normally distributed numbers. They are

$$\Delta x = a_0 y I J \qquad \Delta x^2 = I J (a_0 y^2 + a_2^2 y)$$

From them one immediately has (7).

5 Non-linear effects

The model was constructed for the Gaussian distribution and Gaussian processes. The dynamic regression model was assumed to be linear. However there exist experimental observations in the Framingham pointing to possible non-linear effects: hormesis effects and oscillations in mortality rates [1] as well as deviations in elderly populations (for ages > 75 years).

A resistance or regeneration effect leads to the following generalization of the death process:

$$N^{t+1} = \Big(\eta^t(x) - \mu^t(x)\Big) N,$$

where η is a new variable which is time and risk factor dependent:

$$\eta^{t+1}(x) = F_\eta(\eta^t, x^t), \qquad \mu^{t+1}(x) = F_\mu(\mu^t, x^t).$$

The function η should also be included in equations, to provide the necessary feedback,

$$x_i^{(t+1)} = u_i^0 + \sum_j R_{ij} x_j^{(t)} + \sum_{j,k} C_{ijk} x_j^{(t)} x_k^{(t)} + \\ + \sum_j F_{ij}^{(t)} x_j^{(t)} + \sum_j P_{ij}^{(t)} \eta_j^{(t)} + \epsilon_i$$

This formula is non-linear generalization of one-year model (2). Coefficients C_{ijk} should be estimated from data. The deterministic part is controllable with control accomplished either by adjusting system control parameters or by chaos controlling strategies [21]. F_{ij}^t is a matrix with, generally, time dependent coefficients, determined by controlling strategies. P_{ij} are generally, non-linear functions of x.

6 Summary

In this report general features of life table calculations based on random walks and Kolmogorov-Fokker-Planck equations are discussed. Several schemes for numerical forecasting for different regression formulas (one-year and two-year) and different schemes for filling in of missing data are developed.

Another approach is presented based on the theory of stochastic processes. A scheme for using stochastic processes for numerical predictions is developed and discussed. A life time generator providing simulated data is constructed and is used for numerical and analytical tests.

Possible non-linear effects are also discussed by generalizing the dynamic to include additional factor with non-linear, field dependence and chaotic effects .

References

[1] A. R. D. Stebbing. Growth hormesis: a by-product of control. Health Physics. **52**, no. 5, 543-547 (1987).

[2] K. G. Manton and A. I. Yashin. *Mechanisms of aging and mortality: The search for new paradigms* (Odense University Press, 2000).

[3] F. R. Bayo and J. F. Faber. Mortality rates around age one hundred. Trans.Soc.Act. **35**, 37-59 (1985)

[4] E. A. Lew and L. Garfinkel. Mortality ages 65 and over in a middle-class population. Trans.Soc.Act. **36**, 257-295 (1984).

[5] E. A. Lew and L. Garfinkel. Differences in mortality and longevity by sex, smoking habits and health status. Trans.Soc.Act. **39**, 107-125 (1987).

[6] S. D. Dubey. Some percentile estimators of Weibull parameters. Technometrics. **9**, 119-129 (1967).

[7] B. Rosenberg, G. Kemeny, L. G. Smith, I. D. Skurnick, J. M. Bandurski. The kinetics and thermodynamics of death in multicellular organisms. Mech. Ageing Dev. **2**, 275-293 (1973).

[8] B. L. Strehler, A. S. Mildvan. General theory of mortality and aging. Science. **132**, 14-21 (1960).

[9] L. Hayflick, S. A. Plotkin, T. W. Norton, H. Koprowski. Preparation of poliovirus vaccines in a human fetal diploid cell strain . Amer. J. Hyg. **75**, 240-258 (1962).

[10] J. Smith and L. Martin. Do cells cycle? . Proc. Natl. Acad. Sci. USA. **70**, 1263-1267 (1973).

[11] V. J. Cristofalo, R. G. Allen, R. J. Pignolo, B. G. Martin, J. C. Beck. Relationship between donor age and the replicative life span of human cells in culture: a reevaluation . Proc. Natl. Acad. Sci. USA. **95**, 10614-10619 (1998).

[12] K. S. McCully. *Homocysteine theory of arteriosclerosis: Development and current status.* . In: A. M., Jr Gotto, R. Paoletti (Eds.), Atherosclerosis Reviews (Raven Press, New York, 1983). **11** , pp.157-246.

[13] K. S. McCully. The biomedical significance of homocysteine . Jour. of Sci. Expl. **15**, 5-20 (2001).

[14] M. A. Woodbury and K. G. Manton. A random walk model of human mortality and aging. Theoretical Population Biology. **11**, 37-48 (1977).

[15] M. A. Woodbury, K. G. Manton. A theoretical model of the physiological dynamics of circulatory disease in human populations. Human Biology. **55**, 417-441 (1983).

[16] K. G. Manton and E. Stallard. *Chronic disease modeling: Measurement and evaluation of the risks of chronic disease processes* (Oxford University Press, New York, NY, 1988).

[17] A. I. Yashin and K. G. Manton. Effects of unobserved and partially observed covariate processes on system failure. Statistical Science. **12**, no. 1, 20-34 (1997).

[18] K. G. Manton, E. Stallad, M. A. Woodbury, and J. Dowd. Time-varying covariates in models of human mortality and aging: Multidimensional generalizations of the Gompertz. Journal of Gerontology: Biological Sciences. **49**, B169-B190 (1994).

[19] H. Risken. *The Fokker-Planck equation: Methods of solution and applications*, 2nd edn. (Springer-Verlag, New York, NY, 1989) (1996 printing).

[20] K. G. Manton and E. Stallard. Projecting the future size and health status of the US elderly population. International Journal of Forecasting. **8**, 433-458 (1992).

[21] H. G. Schuster, ed. *Handbook of Chaos Control* (Wiley VCH, Weinheim, 1999).